

---

# One-step Diffusion Models with Bregman Density Ratio Matching

---

**Yuanzhi Zhu<sup>\*1</sup> Eleftherios Tsonis<sup>\*1</sup> Lucas Degeorge<sup>\*1,2,3</sup> Vicky Kalogeiton<sup>1</sup>**

<sup>1</sup>LIX, École Polytechnique, CNRS, IPP <sup>2</sup>LIGM, École Nationale des Ponts et Chaussées, CNRS, IPP <sup>3</sup>AMIAD

## Abstract

Diffusion and flow models achieve high generative quality but remain computationally expensive due to slow multi-step sampling. Distillation methods accelerate them by training fast student generators, yet most existing objectives lack a unified theoretical foundation. In this work, we propose Di-Bregman, a compact framework that formulates diffusion distillation as Bregman divergence-based density-ratio matching. This convex-analytic view connects several existing objectives through a common lens. Experiments on CIFAR-10 and text-to-image generation demonstrate that Di-Bregman achieves improved one-step FID over reverse-KL distillation and maintains high visual fidelity compared to the teacher model. Our results highlight Bregman density-ratio matching as a practical and theoretically-grounded route toward efficient one-step diffusion generation.

## 1 Introduction

Diffusion and flow models [1, 16, 22, 25, 44, 46, 47] have become a cornerstone of generative modeling, attaining state-of-the-art performance across modalities and tasks [5, 7, 9, 10, 10, 26, 37, 39, 50]. Yet their sampling process remains prohibitively slow, often requiring hundreds of network evaluations per sample. This has motivated an active line of research on distillation: training fast student generators that reproduce a pre-trained teacher’s output in one or few steps. Current approaches can be broadly categorized as Ordinary Differential Equation (ODE)-based [13, 23, 41, 45], which learn consistency mappings along the teacher’s probability-flow ODE, and distribution-based [28, 59, 65], which directly match the generator’s output distribution to that of the teacher or data. ODE-based methods enforce sufficient but unnecessary conditions for one-step generation, whereas distribution-based methods relax these constraints and capture a broader solution space. Variational Score Distillation (VSD) [52] and Distribution Matching Distillation (DMD) [59] define objectives based on reverse-Kullback-Leibler (KL) divergence between student and teacher models.  $f$ -distill [55] reframed these methods through the lens of  $f$ -divergences. Despite this progress, a general perspective that explains these objectives in a simple mathematical form remains missing.

We introduce Di-Bregman, a general framework that formulates diffusion distillation as Bregman divergence-based density-ratio matching [32]. The central insight is that aligning the student distribution  $q(x)$  with the teacher  $p(x)$  can be viewed as driving the ratio  $r(x) = \frac{q(x)}{p(x)}$  toward constant one, under a suitable convex function  $h$ . This perspective yields a closed-form gradient (Theorem 3.1) with weighting  $h''(r)r$ . Under this formulation, familiar objectives, such as KL- or MSE-based distillation arise as specific choices of  $h$ . The result is a concise, interpretable expression that connects multiple existing formulations within a single theoretical framework.

Beyond theory, Di-Bregman remains practical. To get the weighting coefficient  $h''(r)r$ , we estimate density ratios through a simple classifier trained to distinguish student samples from real data, enabling efficient training without repeated teacher simulation and allowing optional adversarial

---

<sup>\*</sup> share the same office

refinement. Preliminary results on both unconditional image and text-to-image generation demonstrate that our approach attains improved one-step FID than reverse-KL distillation and maintains visual fidelity comparable to the multi-step teacher models.

In summary, our contributions are:

- We introduce a unified formulation of diffusion distillation based on Bregman density-ratio matching, which yields a closed-form gradient interpretation,
- We propose a practical classifier-based training procedure that effectively instantiates this formulation and validate it on early benchmarks.

## 2 Preliminaries

### 2.1 Variational Score Distillation

Variational Score Distillation (VSD) [52] was introduced to mitigate mode-seeking and over-saturation<sup>2</sup> issues observed when using Score Distillation Sampling (SDS) for 3D asset generation [38]. Importantly, the VSD objective is defined on the *final* samples produced by a generator, rather than on intermediate sampler states. This final-sample focus naturally motivates efforts to distil powerful multi-step pre-trained model into compact few-step or one-step generators via VSD-style objectives; several recent works have followed this route [28, 33, 59].

Concretely, VSD can be viewed as minimizing a time-averaged divergence between the noisy marginal produced by the student generator and the corresponding noisy marginal of a pretrained reference model. Writing  $q_t$  and  $p_t$  for the generator and reference noisy marginals at time  $t$ , respectively, the gradient of a typical score-distillation loss admits the following approximation:

$$\nabla_{\theta} \mathcal{L}_{\text{VSD}} = \mathbb{E}_t [\nabla_{\theta} \text{KL}(q_t \| p_t)] \approx -\mathbb{E}_{t,\epsilon} \left[ w(t) (s_{\phi}(x_t, t) - s_{\psi}(x_t, t)) \frac{dG_{\theta}(\epsilon)}{d\theta} \right], \quad (1)$$

where  $w(t)$  is a scalar weighting function over timesteps, and the noisy state  $x_t$  is obtained by applying the forward diffusion kernel at time  $t$  to the generator output  $G_{\theta}(\epsilon)$  using another independent Gaussian noise.  $s_{\phi}(\cdot, t)$  and  $s_{\psi}(\cdot, t)$  denote the pre-trained score function on reference data and the auxiliary score function on the student-generated data evaluated at timestep  $t$ , respectively. Intuitively, the score difference  $s_{\phi} - s_{\psi}$  provides a learning signal that pushes the student’s generated noisy marginals toward those of the pre-trained teacher model, and backpropagating through  $G_{\theta}$  to update the generator parameters  $\theta$ .

### 2.2 Bregman Divergence for Density Ratio Matching

Given two probability distributions  $p^*(x)$  and  $q^*(x)$ , the goal of *density ratio matching* is to learn a ratio model  $r_{\theta}(x)$  that approximates the true density ratio  $r^*(x) := \frac{q^*(x)}{p^*(x)}$  based on i.i.d. samples drawn from both distributions.

The *Bregman divergence* provides a flexible and theoretically grounded measure for comparing functions such as density ratios. It generalizes the notion of squared Euclidean distance to a broad class of divergences that share similar geometric and convexity properties [2, 49]. Let  $h$  be a differentiable and strictly convex function. The Bregman divergence associated with  $h$  between two functions  $r$  and  $r^*$  is defined as [19, 49]:

$$D_h(r \| r^*) = \int p(x) \left[ h(r(x)) - h(r^*(x)) - h'(r^*(x))(r(x) - r^*(x)) \right] dx. \quad (2)$$

This divergence is positive-definite, which means it is always non-negative and equals zero if and only if  $r(x) = r^*(x)$  almost everywhere with respect to  $p(x)$ , which is the density implicitly defined in  $r(x)$ . Many well-known divergences arise as special cases of the Bregman divergence for specific choices of the convex function  $h$ . For instance, the *squared loss* corresponds to  $h(r) = \frac{1}{2}r^2$ , leading

---

<sup>2</sup>The SDS objective tends to produce solutions corresponding to the mode of the averaged likelihood, leading to mode-seeking behavior. Moreover, a high Classifier Free Guidance (CFG) scale can cause over-saturated and over-smoothed generation results.

to least-squares density ratio estimation [48] and the *KL divergence* corresponds to  $h(r) = r \log r - r$ . More instances can be found in Sec. 2.2. This unifying framework allows density ratio estimation to be interpreted as minimizing a Bregman divergence under different convex function  $h$ , providing a general connection between statistical divergences and convex analysis [2, 6].

Table 1: Examples of different  $h(r)$  in Bregman divergence and the corresponding  $h''(r)r$ . The choices of  $h(r)$  are from [19, 34].

Name	$h(r)$	$h''(r)r$	$h''(e^{-l})e^{-l}$
LR	$r \log r - (1+r) \log(1+r)$	$\frac{1}{1+r}$	$\sigma(l)$
KL	$r \log r - r$	1	1
BE	$-\log r$	$1/r$	$e^l$
LS	$r^2/2$	$r$	$e^{-l}$
SBA	$\frac{r^{1+\lambda}-r}{\lambda(\lambda+1)}$	$r^\lambda$	$e^{-\lambda l}$

### 3 Method

In this section, we introduce a general distillation framework, termed Di-Bregman, which is derived from the Bregman divergence for density ratio matching formulation in Sec. 2.2. The core idea is to align the student distribution  $q(x)$ , induced by a one-step generator  $G_\theta$ , with the teacher distribution  $p(x)$ . Since the student distribution  $q(x)$  is implicitly defined by the generator through the push-forward measure of the prior, i.e.,  $x = G_\theta(\epsilon)$  with  $\epsilon \sim \mathcal{N}(0, I)$ , the distribution  $q(x)$  and its density ratio depend on the generator parameters  $\theta$ . Let  $r(x) = \frac{q(x)}{p(x)}$  denote the density ratio between the student and teacher distributions. Perfect alignment hence corresponds to  $r(x) = 1$  for all  $x$ , which motivates minimizing a divergence between  $r(x)$  and the target ratio 1 in Eq. (2):

$$D_h(r\|1) = \int p(x) \left[ h(r(x)) - h(1) - h'(1)(r(x) - 1) \right] dx. \quad (3)$$

Minimizing this divergence with respect to  $\theta$  encourages the student generator  $G_\theta$  to produce samples whose induced distribution  $q(x)$  matches the teacher distribution  $p(x)$ .

Following prior work on one-step and few-step distillation of diffusion models [28, 33, 59], we can derive the analytical form of the gradient of the Bregman divergence in Eq. (3). The resulting expression corresponds to a weighted variant of the gradient used in the KL-based objective (Eq. (1)), where the weight is a function of the density ratio  $r(x)$ , analogous to the formulation in  $f$ -distill [55].

To further generalize this result as in VSD, we consider the intermediate distributions  $p_t$  and  $q_t$  obtained via the diffusion forward process. This allows the Bregman-based distillation gradient to be evaluated at arbitrary diffusion timesteps. The following theorem formally characterizes the gradient of the Bregman divergence in this general setting.

**Theorem 3.1** (Gradient of Bregman divergence). *Let  $p_t$  be a reference (teacher) marginal density at time  $t$  and let  $q_t = q_{\theta,t}$  be the marginal induced by the generator  $G_\theta$  at time  $t$ . These intermediate densities are obtained via the forward diffusion process. Define the intermediate density ratio  $r_t(x) := \frac{q_{\theta,t}(x)}{p_t(x)}$ . Assume that  $h$  is twice continuously differentiable. Then the gradient of the Bregman divergence  $D_h(r_t\|1) = \mathbb{E}_{p_t}[h(r_t)] - h(1)$  with respect to  $\theta$  admits the following form:*

$$\nabla_\theta D_h(r_t\|1) = -\mathbb{E}_\epsilon \left[ w(t) h''(r_t(x_t)) r_t(x_t) (\nabla_{x_t} \log p_t(x_t) - \nabla_x \log q_{\theta,t}(x_t)) \nabla_\theta G_\theta(\epsilon) \right], \quad (4)$$

where  $w(t)$  is a weight function.

The corresponding proof can be found in Appendix B.

In practice, the density ratio on noisy data,  $r_t(x) = \frac{q_t(x)}{p_t(x)}$ , can be estimated using a classifier trained to distinguish samples from the student generator  $G_\theta$  and those from the teacher model or reference dataset. Under the common assumption that the pre-trained teacher model already captures the data distribution well, it is often both preferable and computationally cheaper to draw real samples



Figure 2: Images generated with only one-step by model trained with Di-Bregman. More images are shown in Appendix E

directly from the dataset rather than repeatedly sampling from the teacher. The discriminator loss is hence:

$$\min_{G_\theta} \max_{D_\eta} \mathbb{E}_{x_{\text{gt}} \sim p_{\text{data}}, t \sim p_{t_{\text{GAN}}}} [\log D_\eta((x_{\text{gt}})_t, t)] + \mathbb{E}_{\epsilon \sim \mathcal{N}(0, 1), t \sim p_{t_{\text{GAN}}}} [\log(1 - D_\eta(G_\theta(\epsilon)_t, t))]. \quad (5)$$

For a discriminator output  $D_\eta = \sigma(l_t(x))$ , where  $l_t(x)$  denotes the classifier logits at noise level  $t$ , the optimal output satisfies  $\sigma(l_t^*(x)) = \frac{p_t(x)}{p_t(x) + q_t(x)}$ . This implies that the density ratio can be recovered as  $r_t(x) = e^{-l_t(x)}$ . We provide common used  $h(r)$  and corresponding  $h''(e^{-l})e^{-l}$  in Sec. 2.2. In this framework, the trained classifier not only provides an estimate of the local density ratio but can also be repurposed as a discriminator for adversarially training the student generator.

Compared to  $f$ -distill [55], our framework places fewer constraints on the convex function, which yields greater flexibility in choosing divergence families for distillation. Recently, Uni-Instruct [51] proposes a unifying view that connects integral  $f$ -divergences [55, 59] and score-based divergences [29, 65]. However, Di-Bregman is complementary to this line of work: it provides a broader class of function  $h$  and recovers many existing objectives as special cases. Together, these formulations offer a more complete picture of distribution-based diffusion distillation.

## 4 Experiments

To evaluate the effectiveness of the proposed method, we conduct experiments on both unconditional image and text-to-image generation tasks. Quantitative results are reported on the CIFAR-10 dataset using an EDM teacher [17], while qualitative results are presented for text-to-image generation with a Stable Diffusion v1.5 [39] teacher. As shown in Fig. 1, when applying the SBA-type Bregman divergence with  $\lambda = 5$ , our method achieves a lower one-step Fréchet Inception Distance (FID) compared to the baseline reverse KL distillation approach. In addition, Fig. 2 illustrates representative one-step samples generated by our distilled text-to-image model, demonstrating high visual quality and fidelity to the text prompts. More experimental details (D), qualitative (E), quantitative (F) results and additional ablations (F) are provided in Appendix.

## 5 Conclusion

We introduced Di-Bregman, a generalized framework for diffusion model distillation grounded in Bregman divergences. Empirically, our method improves one-step generation quality on CIFAR-10 and produces competitive visual results in text-to-image synthesis, demonstrating both its theoretical generality and practical effectiveness.

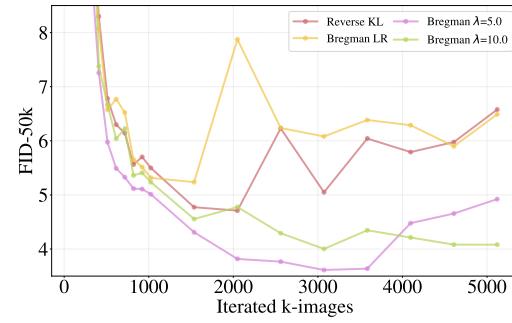


Figure 1: **Evolution of one-step FID against number of iterated images:** Di-Bregman achieves a lower one-step FID.

---

**Algorithm 1** Di-Bregman Distillation

---

**Require:** Pre-trained teacher model  $\phi$ , auxiliary model  $\psi$ , discriminator heads  $\eta$ , condition dataset  $\mathcal{D}_c$ , ground truth dataset  $\mathcal{D}_d$ , loss weights  $w_{\text{GAN}}$  (optional)

1:  $\theta \leftarrow \text{copyWeights}(\phi)$ ,  $\psi \leftarrow \text{copyWeights}(\phi)$  // initialize models

2: **repeat**

3:   **### Generate one-step image samples  $x_\theta$**

4:   Sample  $\epsilon \sim \mathcal{N}(0, 1)$ ,  $c \sim \mathcal{D}_c$

5:    $x_\theta = G_\theta(\epsilon, c)$

6:   **### Update generator  $\theta$**

7:   Sample  $t \sim \mathcal{U}[0, 1]$ ,  $x_t \sim q_{t|0}(x_t | x_\theta)$  // Forward

8:   Calculate true and auxiliary score  $s_\phi(x_t, c)$  and  $s_\psi(x_t, c)$

9:   Calculate the density ratio  $r_t$  using discriminator head logit output  $r_t(x) = e^{-l_t(x)}$

10:   **# calculate Di-Bregman loss gradient**

11:    $\nabla_\theta \mathcal{L}_{\text{Di-Bregman}}(\theta) \leftarrow -\mathbb{E}_\epsilon \left[ h''(r_t(x_t)) r_t(x_t) (s_\phi(x_t, c) - s_\psi(x_t, c)) \nabla_\theta G_\theta(\epsilon) \right].$

12:   **# calculate GAN loss (optional)**

13:    $\mathcal{L}_{\text{GAN}}(\theta) \leftarrow \mathbb{E}_{\epsilon, t}[-\log(D_\eta(x_t, t))].$

14:   **# calculate total loss and update**

15:   Update  $\theta$  using gradient of  $\mathcal{L}_{\text{gen}}(\theta) = \mathcal{L}_{\text{Di-Bregman}}(\theta) + w_{\text{GAN}} \mathcal{L}_{\text{GAN}}(\theta)$

16:   **### Update auxiliary model  $\psi$**

17:   Sample  $t' \sim \mathcal{U}[0, 1]$ ,  $x_{t'} \sim q_{t|0}(x_{t'} | x_\theta(x_{\text{init}}, c))$

18:   Update  $\psi$  with standard denoising score match loss to learn  $x_\theta$

19:   **### Update discriminator  $\eta$**

20:   Sample  $t'' \sim \mathcal{U}[0, 0.95]$ , calculate  $x_{t''} \sim q_{t|0}(x_{t''} | x_\theta)$

21:   Sample real data  $(x_{\text{gt}}, c) \sim \mathcal{D}_d$ , calculate  $(x_{\text{gt}})_{t''}$

22:   Update  $\eta$  with GAN objective (Eq. (5))

23: **until** convergence

24: **Return** one-step generator  $\theta$

---

## References

- [1] Michael S Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *arXiv preprint arXiv:2303.08797*, 2023.
- [2] Arindam Banerjee, Srujana Merugu, Inderjit S Dhillon, and Joydeep Ghosh. Clustering with bregman divergences. *Journal of machine learning research*, 6(Oct):1705–1749, 2005.
- [3] Nicholas M Boffi, Michael S Albergo, and Eric Vanden-Eijnden. How to build a consistency model: Learning flow maps via self-distillation. *arXiv preprint arXiv:2505.18825*, 2025.
- [4] Nicholas Matthew Boffi, Michael Samuel Albergo, and Eric Vanden-Eijnden. Flow map matching with stochastic interpolants: A mathematical framework for consistency models. *Transactions on Machine Learning Research*, 2025.
- [5] Luc Boudier, Loris Manganelli, Eleftherios Tsonis, Nicolas Dufour, and Vicky Kalogeiton. Training-free synthetic data generation with dual ip-adapter guidance. In *British Machine Vision Conference (BMVC)*, 2025.
- [6] Lev M Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR computational mathematics and mathematical physics*, 7(3):200–217, 1967.
- [7] Robin Courant, Xi Wang, David Loiseaux, Marc Christie, and Vicky Kalogeiton. Pulp motion: Framing-aware multimodal camera and human motion generation. *arXiv preprint arXiv:2510.05097*, 2025.
- [8] Trung Dao, Thuan Hoang Nguyen, Thanh Le, Duc Vu, Khoi Nguyen, Cuong Pham, and Anh Tran. Swiftbrush v2: Make your one-step diffusion model better than its teacher. In *European Conference on Computer Vision*, pages 176–192. Springer, 2025.
- [9] Lucas Degeorge, Arijit Ghosh, Nicolas Dufour, David Picard, and Vicky Kalogeiton. How far can we go with imagenet for text-to-image generation? *arXiv*, 2025.
- [10] Nicolas Dufour, Victor Besnier, Vicky Kalogeiton, and David Picard. Don’t drop your samples! coherence-aware training benefits conditional diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6264–6273, 2024.
- [11] Nicolas Dufour, Lucas Degeorge, Arijit Ghosh, Vicky Kalogeiton, and David Picard. Miro: Multi-reward conditioned pretraining improves t2i quality and efficiency. *arXiv preprint arXiv:2510.25897*, 2025.
- [12] Kevin Frans, Danijar Hafner, Sergey Levine, and Pieter Abbeel. One step diffusion via shortcut models. *arXiv preprint arXiv:2410.12557*, 2024.
- [13] Zhengyang Geng, Mingyang Deng, Xingjian Bai, J Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling. *arXiv preprint arXiv:2505.13447*, 2025.
- [14] Jiatao Gu, Shuangfei Zhai, Yizhe Zhang, Lingjie Liu, and Joshua M Susskind. Boot: Data-free distillation of denoising diffusion models with bootstrapping. In *ICML 2023 Workshop on Structured Probabilistic Inference {\&} Generative Modeling*, 2023.
- [15] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017.
- [16] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33:6840–6851, 2020.
- [17] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in neural information processing systems*, 35:26565–26577, 2022.

- [18] Dongjun Kim, Chieh-Hsin Lai, Wei-Hsiang Liao, Naoki Murata, Yuhta Takida, Toshimitsu Uesaka, Yutong He, Yuki Mitsuji, and Stefano Ermon. Consistency trajectory models: Learning probability flow ode trajectory of diffusion. *arXiv preprint arXiv:2310.02279*, 2023.
- [19] Yeongmin Kim, Heesun Bae, Byeonghu Na, and Il-Chul Moon. Preference optimization by estimating the ratio of the data distribution. *arXiv preprint arXiv:2505.19601*, 2025.
- [20] Alex Krizhevsky and Geoffrey Hinton. Learning multiple layers of features from tiny images. Technical Report 0, University of Toronto, Toronto, Ontario, 2009.
- [21] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Computer vision–ECCV 2014: 13th European conference, zurich, Switzerland, September 6–12, 2014, proceedings, part v 13*, pages 740–755. Springer, 2014.
- [22] Yaron Lipman, Ricky TQ Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2022.
- [23] Qiang Liu. Rectified flow: A marginal preserving approach to optimal transport. *arXiv preprint arXiv:2209.14577*, 2022.
- [24] Qiang Liu. Icml tutorial on the blessing of flow. *International conference on machine learning*, 2025.
- [25] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.
- [26] Shilin Lu, Zilan Wang, Leyang Li, Yanzhu Liu, and Adams Wai-Kin Kong. Mace: Mass concept erasure in diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6430–6440, 2024.
- [27] Eric Luhman and Troy Luhman. Knowledge distillation in iterative generative models for improved sampling speed. *arXiv preprint arXiv:2101.02388*, 2021.
- [28] Weijian Luo, Tianyang Hu, Shifeng Zhang, Jiacheng Sun, Zhenguo Li, and Zhihua Zhang. Diff-instruct: A universal approach for transferring knowledge from pre-trained diffusion models. *Advances in Neural Information Processing Systems*, 36:76525–76546, 2023.
- [29] Weijian Luo, Zemin Huang, Zhengyang Geng, J Zico Kolter, and Guo-jun Qi. One-step diffusion distillation through score implicit matching. *Advances in Neural Information Processing Systems*, 37:115377–115408, 2024.
- [30] Weijian Luo, Zemin Huang, Zhengyang Geng, J Zico Kolter, and Guo-jun Qi. One-step diffusion distillation through score implicit matching. *Advances in Neural Information Processing Systems*, 37:115377–115408, 2025.
- [31] Chenlin Meng, Robin Rombach, Ruiqi Gao, Diederik Kingma, Stefano Ermon, Jonathan Ho, and Tim Salimans. On distillation of guided diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14297–14306, 2023.
- [32] Shakir Mohamed and Balaji Lakshminarayanan. Learning in implicit generative models. *arXiv preprint arXiv:1610.03483*, 2016.
- [33] Thuan Hoang Nguyen and Anh Tran. Swiftbrush: One-step text-to-image diffusion model with variational score distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7807–7816, 2024.
- [34] Frank Nielsen and Richard Nock. Sided and symmetrized bregman centroids. *IEEE transactions on Information Theory*, 55(6):2882–2904, 2009.
- [35] Jean Pachebat, Giovanni Conforti, Alain Durmus, and Yazid Janati. Iterative tilting for diffusion fine-tuning. *arXiv preprint arXiv:2512.03234*, 2025.

- [36] Gaurav Parmar, Richard Zhang, and Jun-Yan Zhu. On aliased resizing and surprising subtleties in gan evaluation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11410–11420, 2022.
- [37] William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4195–4205, 2023.
- [38] Ben Poole, Ajay Jain, Jonathan T Barron, and Ben Mildenhall. Dreamfusion: Text-to-3d using 2d diffusion. *arXiv preprint arXiv:2209.14988*, 2022.
- [39] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695, 2022.
- [40] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.
- [41] Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. *arXiv preprint arXiv:2202.00512*, 2022.
- [42] Tim Salimans, Thomas Mensink, Jonathan Heek, and Emiel Hoogeboom. Multistep distillation of diffusion models via moment matching. *Advances in Neural Information Processing Systems*, 37:36046–36070, 2025.
- [43] Christoph Schuhmann, Romain Beaumont, Richard Vencu, Cade Gordon, Ross Wightman, Mehdi Cherti, Theo Coombes, Aarush Katta, Clayton Mullis, Mitchell Wortsman, et al. Laion-5b: An open large-scale dataset for training next generation image-text models. *neurips*, 2022.
- [44] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep unsupervised learning using nonequilibrium thermodynamics. In *International Conference on Machine Learning*, pages 2256–2265. PMLR, 2015.
- [45] Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *International conference on machine learning*. PMLR, 2023.
- [46] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. *Advances in Neural Information Processing Systems*, 32, 2019.
- [47] Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2020.
- [48] Masashi Sugiyama, Shinichi Nakajima, Hisashi Kashima, Paul von Bünau, and Motoaki Kawanabe. Direct importance estimation with model selection and its application to covariate shift adaptation. *Advances in Neural Information Processing Systems*, 20, 2008.
- [49] Masashi Sugiyama, Taiji Suzuki, and Takafumi Kanamori. Density-ratio matching under the bregman divergence: a unified framework of density-ratio estimation. *Annals of the Institute of Statistical Mathematics*, 64(5):1009–1044, 2012.
- [50] Xi Wang, Robin Courant, Marc Christie, and Vicky Kalogeiton. Akira: Augmentation kit on rays for optical video generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 2609–2619, 2025.
- [51] Yifei Wang, Weimin Bai, Colin Zhang, Debing Zhang, Weijian Luo, and He Sun. Uni-instruct: One-step diffusion model through unified diffusion divergence instruction. *arXiv preprint arXiv:2505.20755*, 2025.
- [52] Zhengyi Wang, Cheng Lu, Yikai Wang, Fan Bao, Chongxuan Li, Hang Su, and Jun Zhu. Proflificdreamer: High-fidelity and diverse text-to-3d generation with variational score distillation. *Advances in Neural Information Processing Systems*, 36, 2024.

- [53] Sirui Xie, Zhisheng Xiao, Diederik P Kingma, Tingbo Hou, Ying Nian Wu, Kevin Patrick Murphy, Tim Salimans, Ben Poole, and Ruiqi Gao. Em distillation for one-step diffusion models. *arXiv preprint arXiv:2405.16852*, 2024.
- [54] Yanwu Xu, Yang Zhao, Zhisheng Xiao, and Tingbo Hou. Ufogen: You forward once large scale text-to-image generation via diffusion gans. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8196–8206, 2024.
- [55] Yilun Xu, Weili Nie, and Arash Vahdat. One-step diffusion models with  $f$ -divergence distribution matching. *arXiv preprint arXiv:2502.15681*, 2025.
- [56] Hanshu Yan, Xingchao Liu, Jiachun Pan, Jun Hao Liew, Qiang Liu, and Jiashi Feng. Perflow: Piecewise rectified flow as universal plug-and-play accelerator. *arXiv preprint arXiv:2405.07510*, 2024.
- [57] Mingxuan Yi, Zhanxing Zhu, and Song Liu. Monoflow: Rethinking divergence gans via the perspective of wasserstein gradient flows. In *International Conference on Machine Learning*, pages 39984–40000. PMLR, 2023.
- [58] Tianwei Yin, Michaël Gharbi, Taesung Park, Richard Zhang, Eli Shechtman, Fredo Durand, and Bill Freeman. Improved distribution matching distillation for fast image synthesis. *Advances in neural information processing systems*, 37:47455–47487, 2024.
- [59] Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Fredo Durand, William T Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6613–6623, 2024.
- [60] Mingtian Zhang, Jiajun He, Wenlin Chen, Zijing Ou, José Miguel Hernández-Lobato, Bernhard Schölkopf, and David Barber. Towards training one-step diffusion models without distillation. *arXiv preprint arXiv:2502.08005*, 2025.
- [61] Haoyang Zheng, Xinyang Liu, Cindy Xiangrui Kong, Nan Jiang, Zheyuan Hu, Weijian Luo, Wei Deng, and Guang Lin. Ultra-fast language generation via discrete diffusion divergence instruct. *arXiv preprint arXiv:2509.25035*, 2025.
- [62] Mingyuan Zhou, Yi Gu, and Zhendong Wang. Few-step diffusion via score identity distillation. *arXiv preprint arXiv:2505.12674*, 2025.
- [63] Mingyuan Zhou, Zhendong Wang, Huangjie Zheng, and Hai Huang. Long and short guidance in score identity distillation for one-step text-to-image generation. *arXiv preprint arXiv:2406.01561*, 2024.
- [64] Mingyuan Zhou, Huangjie Zheng, Yi Gu, Zhendong Wang, and Hai Huang. Adversarial score identity distillation: Rapidly surpassing the teacher in one step. *arXiv preprint arXiv:2410.14919*, 2024.
- [65] Mingyuan Zhou, Huangjie Zheng, Zhendong Wang, Mingzhang Yin, and Hai Huang. Score identity distillation: Exponentially fast distillation of pretrained diffusion models for one-step generation. In *Forty-first International Conference on Machine Learning*, 2024.
- [66] Yuanzhi Zhu. On the connection between dmd and gan for diffusion distillation. Blog post, Nov 2025.
- [67] Yuanzhi Zhu, Xingchao Liu, and Qiang Liu. Slimflow: Training smaller one-step diffusion models with rectified flow. In *European Conference on Computer Vision*, pages 342–359. Springer, 2025.
- [68] Yuanzhi Zhu, Xi Wang, Stéphane Lathuilière, and Vicky Kalogeiton. Di[M]o: Distilling masked diffusion models into one-step generator. *arXiv preprint arXiv:2503.15457*, 2025.
- [69] Yuanzhi Zhu, Xi Wang, Stéphane Lathuilière, and Vicky Kalogeiton. Soft-di [m] o: Improving one-step discrete image generation with soft embeddings. *arXiv preprint arXiv:2509.22925*, 2025.

## A Limitations and Future Works.

This work primarily presents preliminary results. In future studies, we plan to extend our approach to a wider range of teacher models and conduct comprehensive comparisons with state-of-the-art methods. Moreover, while our current experiment only use the classifier in Eq. (4), we aim to incorporate adversarial training based on it to further enhance the performance of one-step generation.

## B Derivation

**Theorem B.0** (Gradient of Bregman divergence). *Let  $p_t$  be a reference (teacher) marginal density at time  $t$  and let  $q_t = q_{\theta,t}$  be the marginal induced by the generator  $G_\theta$  at time  $t$ . These intermediate densities are obtained via the forward diffusion process. Define the intermediate density ratio  $r_t(x) := \frac{q_{\theta,t}(x)}{p_t(x)}$ . Assume that  $h$  is twice continuously differentiable. Then the gradient of the Bregman divergence  $D_h(r_t \| 1) = \mathbb{E}_{p_t}[h(r_t)] - h(1)$  with respect to  $\theta$  admits the following form:*

$$\nabla_\theta D_h(r_t \| 1) = -\mathbb{E}_\epsilon \left[ w(t) h''(r_t(x_t)) r_t(x_t) (\nabla_{x_t} \log p_t(x_t) - \nabla_x \log q_{\theta,t}(x_t)) \nabla_\theta G_\theta(\epsilon) \right]. \quad (6)$$

where  $w(t)$  is a weight function.

*Proof.* Recall that

$$D_h(r_t \| 1) = \int p_t(x) \left[ h(r_t(x)) - h(1) - h'(1)(r_t(x) - 1) \right] dx, \quad (7)$$

where  $r_t(x) = q_{\theta,t}(x)/p_t(x)$  and  $p_t$  do not depend on  $\theta$ . Differentiating under the integral sign and using that

$$\int p_t(x) \nabla_\theta r_t(x) dx = \nabla_\theta \int q_{\theta,t}(x) dx = \nabla_\theta 1 = 0, \quad (8)$$

we obtain

$$\nabla_\theta D_h(r_t \| 1) = \int p_t(x) h'(r_t(x)) \nabla_\theta r_t(x) dx = \int h'(r_t(x)) \nabla_\theta q_{\theta,t}(x) dx. \quad (9)$$

Next, we express  $q_{\theta,t}$  as the pushforward of a base noise  $\epsilon \sim p(\epsilon)$  through the generator at time  $t$ ,  $x_t = F(G_\theta(\epsilon), z)$  with fixed forward process  $F(x, z) = \alpha_t x + \sigma z$  and  $z \sim \mathcal{N}(0, I)$ :

$$q_{\theta,t}(x) = \int p(\epsilon) \delta(x - x_t) d\epsilon, \quad (10)$$

where  $\delta$  is the Dirac delta. Differentiating equation 10 w.r.t.  $\theta$  and using the chain rule for distributions yields

$$\nabla_\theta q_{\theta,t}(x) = \int p(\epsilon) \nabla_\theta \delta(x - x_t) d\epsilon = -w(t) \int p(\epsilon) \nabla_x \delta(x - x_t) \nabla_\theta G_\theta(\epsilon) d\epsilon. \quad (11)$$

Substituting equation 11 into equation 9 gives

$$\nabla_\theta D_h(r_t \| 1) = - \int w(t) h'(r_t(x)) \left[ \int p(\epsilon) \nabla_x \delta(x - x_t) \nabla_\theta G_\theta(\epsilon) d\epsilon \right] dx \quad (12)$$

$$= - \int p(\epsilon) \left[ w(t) \int h'(r_t(x)) \nabla_x \delta(x - x_t) dx \right] \nabla_\theta G_\theta(\epsilon) d\epsilon. \quad (13)$$

We now integrate by parts in  $x$  (assuming boundary terms vanish):

$$\int h'(r_t(x)) \nabla_x \delta(x - x_t) dx = - \int \delta(x - x_t) \nabla_x h'(r_t(x)) dx. \quad (14)$$

Hence

$$\begin{aligned} \nabla_\theta D_h(r_t \| 1) &= \int p(\epsilon) \left[ \int w(t) \delta(x - x_t) \nabla_x h'(r_t(x)) dx \right] \nabla_\theta G_\theta(\epsilon) d\epsilon \\ &= \mathbb{E}_\epsilon [w(t) \nabla_x h'(r_t(x_t)) \nabla_\theta G_\theta(\epsilon)]. \end{aligned} \quad (15)$$

Apply the chain rule  $\nabla_x h'(r_t) = h''(r_t) \nabla_x r_t$  and  $\nabla_x r_t = r_t (\nabla_x \log q_{\theta,t} - \nabla_x \log p_t)$  yields Eq. (4) stated in the theorem. An alternative proof can be constructed following the approach of [57].  $\square$

## C Related Works: Diffusion Distillation

Distillation methods for accelerating diffusion and flow models fall into two broad families. *ODE-based* distillation exploits the teacher’s Probability Flow ODE (PF-ODE) to derive regression-style objectives for a student model [3, 4, 12, 13, 14, 18, 23, 27, 31, 41, 45, 56, 67]. These approaches frame distillation as learning an ODE-consistent mapping, often enabling stable one- or few-step samplers which preserves the coupling induced by teacher models’ PF-ODE. By contrast, *distribution-based* methods align the student generator’s output distribution with the teacher’s multi-step sampling distribution (or with a specified data distribution) without relying on an explicit PF-ODE. This class covers divergence- and adversarial-style matching techniques [8, 28, 30, 33, 42, 51, 52, 53, 54, 58, 59, 61, 62, 63, 64, 65, 68, 69]. Compared to distribution-based methods, ODE-based formulations optimize more indirect objectives that enforce consistency with an underlying continuous-time dynamics. These ODE constraints are *sufficient but not necessary* for correct one-step generation. Consequently, ODE-based methods are more restrictive, while distribution-based formulations directly match the target distribution and thus allow a broader family of solutions and greater modeling flexibility. In *f-distil*, [55] extend the VSD framework from reverse Kullback–Leibler (KL) divergence to more general *f*-divergence and use discriminator to estimate the density ratio. A notable feature of many distribution-based methods is that they match not only the final data distribution but also the intermediate noisy-data distributions encountered during sampling; this property has also been referred to as *Interpolation Distillation* [24].

## D Experimental Setup

**Datasets and Pre-trained Teacher Models.** Our experiments to demonstrate the effectiveness of Di-Bregman are performed on the CIFAR-10 [20]  $32 \times 32$  for unconditional generation and on the LAION [43] and COCO [21] datasets for text-to-image generation. The pre-trained teacher models are adopted from the official checkpoints from previous works, EDM [17], and Stable Diffusion v1.5 [39].

**Implementation Details.** All experiments are conducted on a single NVIDIA H100 GPU. For CIFAR-10 ( $32 \times 32$ ) experiments, we adopt the U-Net architecture of NCSN++ [47]. The implementation is based on the SiDA framework [64], where the discriminator is built upon the auxiliary model encoder, and the mean feature vector is used as the predicted discriminator logits. For the text-to-image experiment, we use Stable-Diffusion v1.5 [39], a 900M-paramenter U-net-based model, trained on LAION [43] and distilled at  $512 \times 512$  resolution. All results presented in the paper are one-step generated using our distilled generator.

**Evaluation Metrics.** The metrics we use for quantitative results on CIFAR-10 are Fréchet Inception Distance (FID) [15] and Inception Score (IS) [40]. In our experiments, FID is computed with 50,000 generated samples compared against the training set using Clean-FID [36], while IS is calculated from the same generated images based on their Inception features.

## E Additional qualitative results

In Fig. 3, we provide additional qualitative comparisons, where our one-step student produces visually coherent and faithful samples, closely matching the teacher output across diverse prompts. Additional uncurated CIFAR-10 samples from our Di-Bregman model are shown in Fig. 6, demonstrating diverse one-step generation, with an FID of 3.61.

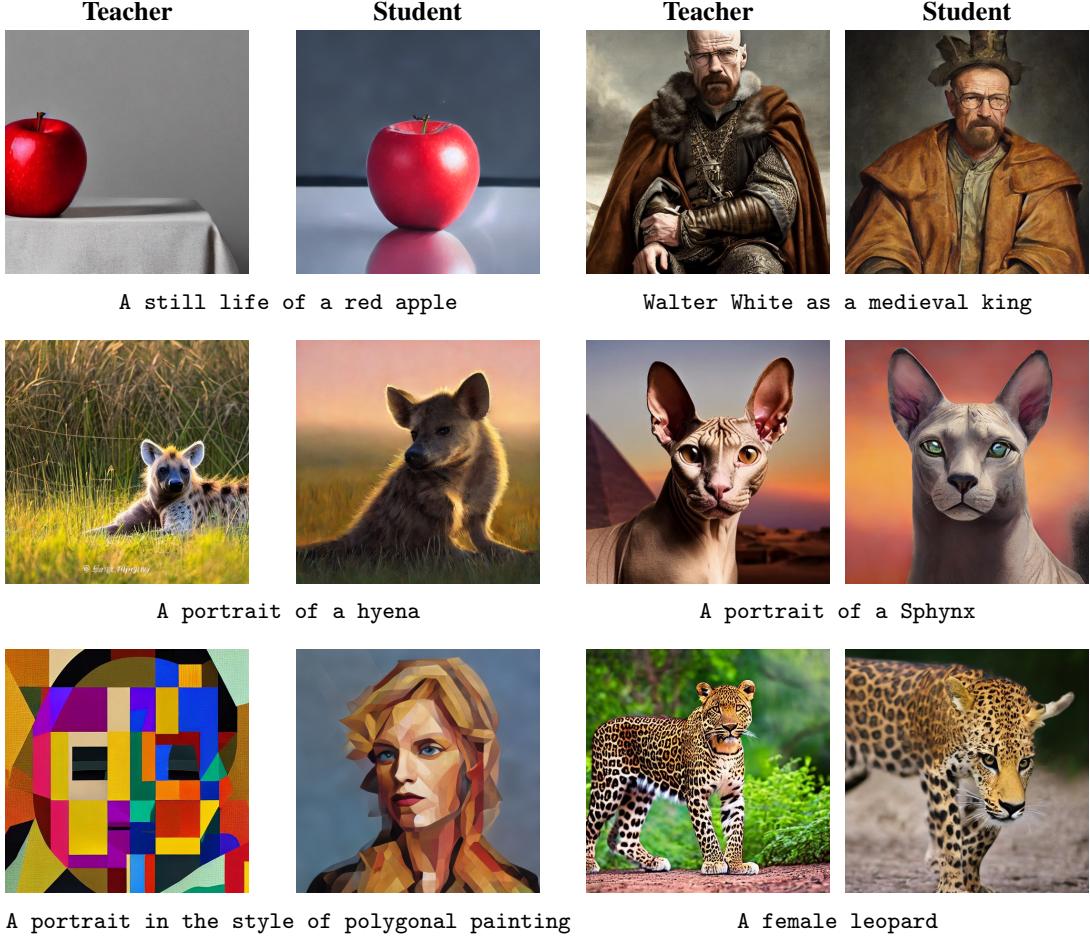


Figure 3: **Qualitative comparison at  $512 \times 512$ : Teacher 50 NFEs (first, third columns) vs Student 1 NFE (second, last columns) for six prompts (left and right blocks per row).** The teacher is the Stable Diffusion v1.5 [39] model.

## F Additional quantitative results

In this section, we provide some additional quantitative results from our one-step models.

Figures 4 and 5 show the evolution of one-step FID and IS, respectively, across different values of the Bregman parameter  $\lambda$ . We observe that Di-Bregman consistently improves over the reverse KL baseline for several  $\lambda$  configurations. In particular, settings such as  $\lambda = 3.0$ ,  $\lambda = 5.0$ , and  $\lambda = 10.0$  yield the lowest one-step FID (Fig. 4) and the highest one-step IS (Fig. 5), confirming the robustness of Di-Bregman across a range of divergence parameters. Lower  $\lambda$  values (e.g.,  $\lambda \leq 1.0$ ) tend to perform closer to the baseline, while negative  $\lambda = -1.0$  underperforms. These results demonstrate that Di-Bregman offers consistent improvements in sample quality metrics over the reverse KL distillation method.

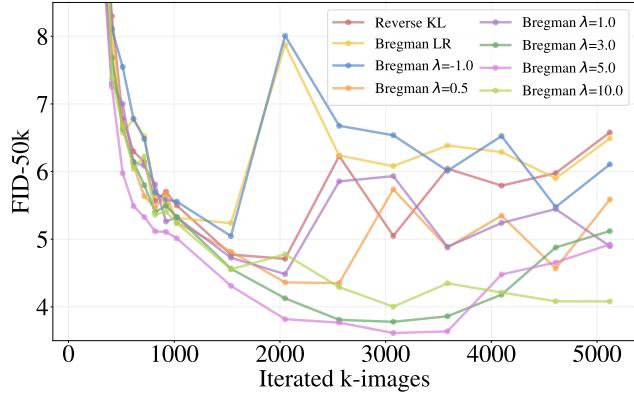


Figure 4: **Evolution of one-step FID against number of iterated images.** For  $\lambda = 3.0, \lambda = 5.0$  or  $\lambda = 10.0$  Di-Bregman achieves a lower one-step FID.

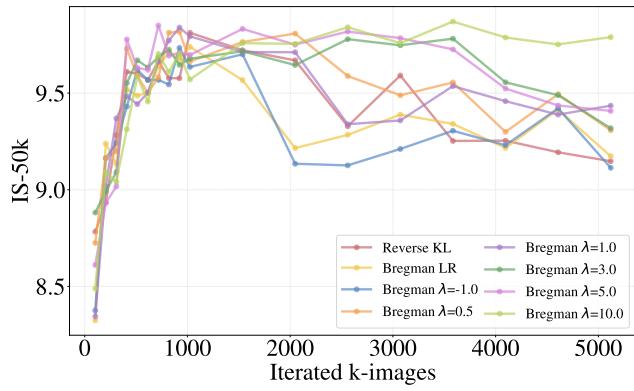


Figure 5: **Evolution of one-step IS against number of iterated images.** For  $\lambda = 3.0, \lambda = 5.0$  or  $\lambda = 10.0$  Di-Bregman achieves a higher one-step IS.



Figure 6: Uncurated samples from unconditional CIFAR-10  $32 \times 32$  using Di-Bregman with single step generation (FID=3.61).

## G General Divergence with a Fixed Reward-Tilted Target Ratio

We begin by defining a reward-tilted target distribution at the clean data level  $x_0$  [35]. Given a base distribution  $p(x_0)$  and an unnormalized reward function  $R(x_0) \geq 0$ , the corresponding reward-tilted clean distribution is

$$p^*(x_0) \propto R(x_0) p(x_0) = \frac{R(x_0)}{Z_0} p(x_0), \quad Z_0 := \mathbb{E}_p[R(x_0)]. \quad (16)$$

To obtain the corresponding target distribution at diffusion time  $t$ , we push  $p^*(x_0)$  through the fixed forward noising kernel  $q_t(x_t | x_0)$ :

$$p_t^*(x_t) = \int q_t(x_t | x_0) p^*(x_0) dx_0 = \frac{1}{Z_0} \int R(x_0) p(x_0) q_t(x_t | x_0) dx_0. \quad (17)$$

Using the marginal identity

$$p_t(x_t) = \int p(x_0) q_t(x_t | x_0) dx_0, \quad (18)$$

the tilted marginal  $p_t^*$  admits the ratio representation

$$p_t^*(x_t) = \frac{R_t(x_t)}{Z_0} p_t(x_t), \quad R_t(x_t) := \mathbb{E}[R(x_0) | x_t] = \frac{\int R(x_0) p(x_0) q_t(x_t | x_0) dx_0}{p_t(x_t)}. \quad (19)$$

Thus the unnormalized intermediate reward  $R_t(x_t)$  is precisely the conditional expectation of the clean reward.

### G.1 Intermediate Unnormalized Reward

The function  $R_t$  is the unique quantity for which

$$p_t^*(x_t) \propto R_t(x_t) p_t(x_t) = \int q_t(x_t | x_0) R(x_0) p(x_0) dx_0. \quad (20)$$

Equivalently, by Bayes' rule,

$$R_t(x_t) = \mathbb{E}[R(x_0) | x_t] = \int R(x_0) p(x_0 | x_t) dx_0. \quad (21)$$

**Taylor Expansion Around the Posterior Mean.** Let

$$\mu_{0,t}(x_t) := \mathbb{E}[x_0 | x_t], \quad \Sigma_{0,t}(x_t) := \text{Cov}(x_0 | x_t). \quad (22)$$

Expanding  $R(x_0)$  around  $\mu_{0,t}$  gives

$$R(x_0) \approx R(\mu_{0,t}) + \nabla R(\mu_{0,t})^\top (x_0 - \mu_{0,t}) + \frac{1}{2} (x_0 - \mu_{0,t})^\top \nabla^2 R(\mu_{0,t}) (x_0 - \mu_{0,t}). \quad (23)$$

Taking  $\mathbb{E}[\cdot | x_t]$  yields

$$\mathbb{E}[R(x_0) | x_t] \approx R(\mu_{0,t}) + \frac{1}{2} \text{Tr}(\nabla^2 R(\mu_{0,t}) \Sigma_{0,t}). \quad (24)$$

**First-Order Approximation.** When the posterior variance is small (e.g. low noise or highly informative forward kernels), the second-order term is negligible:

$$R_t(x_t) = \mathbb{E}[R(x_0) | x_t] \approx R(\mu_{0,t}(x_t)). \quad (25)$$

This approximation is widely used in classifier guidance, reward-guided diffusion, and flow-based preference shaping [24].

## G.2 Normalizing Constant

The intermediate normalizer is

$$Z_t = \mathbb{E}_{x_t \sim p_t} [R_t(x_t)] = \int p_t(x_t) R_t(x_t) dx_t. \quad (26)$$

Substituting the definition of  $R_t$  and applying the law of total expectation,

$$Z_t = \mathbb{E}_{x_t} [\mathbb{E}[R(x_0) | x_t]] = \mathbb{E}_{x_0 \sim p} [R(x_0)] =: Z_0. \quad (27)$$

Hence the normalizing constant is *independent of t*, and we use  $Z_0$  for this normalizing constant across different  $t$ :

$$Z_t = Z_0 = \mathbb{E}_{x_0 \sim p} [R(x_0)] \quad \text{for all } t. \quad (28)$$

Defining the normalized intermediate ratio

$$\tilde{R}_t(x_t) := \frac{R_t(x_t)}{Z_0} = \frac{\mathbb{E}[R(x_0) | x_t]}{\mathbb{E}_p[R(x_0)]}, \quad (29)$$

and using the first-order approximation,

$$\tilde{R}_t(x_t) \approx \frac{R(\mu_{0,t}(x_t))}{\mathbb{E}_p[R(x_0)]}. \quad (30)$$

In practice, assuming new samples arise from the same underlying clean distribution  $p$ , the normalizer may be estimated once from an empirical dataset [11]:

$$\hat{Z} = \frac{1}{N} \sum_{i=1}^N R(x_0^{(i)}), \quad x_0^{(i)} \sim p(x_0), \quad (31)$$

and reused for all  $t$ .

## G.3 Gradient of the General Bregman Divergence

The Bregman divergence between the model ratio  $r_t$  and target ratio  $r_t^*$  is

$$D_h(r_t \| r_t^*) = \int p_t(x) \left( h(r_t(x)) - h(r_t^*(x)) - h'(r_t^*(x)) (r_t(x) - r_t^*(x)) \right) dx. \quad (32)$$

Differentiating w.r.t.  $\theta$ ,

$$\nabla_\theta D_h(r_t \| r_t^*) = \int p_t(x) h'(r_t(x)) \nabla_\theta r_t(x) dx - \int p_t(x) h'(r_t^*(x)) \nabla_\theta r_t(x) dx. \quad (33)$$

When  $r_t^* \equiv 1$ , the second term vanishes, recovering the default Di-Bregman gradient. Under reward tilting, however,  $r_t^*(x) = \frac{R_t(x_t)}{Z_0}$  is no longer constant and this correction term must be preserved.

**Final Gradient Form.** Using

$$\nabla_\theta r_t(x_t) = r_t(x_t) (\nabla_{x_t} \log p_t(x_t) - \nabla_{x_t} \log q_{\theta,t}(x_t)) \nabla_\theta G_\theta(\epsilon), \quad x_t = G_\theta(\epsilon), \quad (34)$$

and the chain rule for  $r_t^*(x_t)$ , we obtain

$$\nabla_\theta D_h(r_t \| r_t^*) = -\mathbb{E}_\epsilon \left[ h''(r_t(x_t)) r_t(x_t) (\nabla_{x_t} \log p_t(x_t) - \nabla_{x_t} \log q_{\theta,t}(x_t)) \nabla_\theta G_\theta(\epsilon) - h''(r_t^*(x_t)) \nabla_\theta r_t^*(x_t) \right]. \quad (35)$$

The first term matches default Di-Bregman training, while the second term encodes a *reward-induced correction* arising from the non-constant target ratio.

**Interpretation as Reward-Tilted Distillation.** The reward-tilted target distribution at time  $t$  satisfies

$$p_t^*(x_t) \propto \frac{R_t(x_t)}{Z_0} p_t(x_t), \quad r_t^*(x_t) = \tilde{R}_t(x_t) = \frac{R_t(x_t)}{Z_0}. \quad (36)$$

Thus Eqs. (33) and (35) provide a principled mechanism for distilling a reward-tilted target into a one-step generator. The normalizer  $Z_0$  acts only as a global scale and does not affect the gradient structure, meaning any unnormalized reward  $R$  can be used as long as it is properly propagated to time  $t$ .

This formulation unifies classical DMD ( $r_t^* \equiv 1$ ), reward-based refinement, and discriminator-driven density-ratio matching into a single framework for reward-tilted generative modeling.

## H Distillation without a Teacher Model

Observe that in our original formulation the teacher model appears *only* through the true score appearing in the score difference term. Using the identity  $\nabla_x r_t = r_t(\nabla_x \log q_{\theta,t} - \nabla_x \log p_t)$ , we can replace this score difference with an estimate derived solely from a discriminator. Specifically, given a discriminator  $D_\eta$ , the density ratio is recovered via  $r_t = \frac{1-D_\eta}{D_\eta}$ . Plugging this into the Bregman objective yields the following gradient:

$$\nabla_\theta D_h(r_t \| 1) = -\mathbb{E}_\epsilon \left[ w(t) h''(r_t(x_t)) \nabla_{x_t} r_t(x_t) \nabla_\theta G_\theta(\epsilon) \right]. \quad (37)$$

This update has the same structural form as the GAN generator gradient. Hence, our Di-Bregman objective naturally recovers GAN training rules when a teacher score is unavailable [60, 66].